SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Mixson, James A

(ii) TITLE OF INVENTION: CARRIER:NUCLEIC ACIDS COMPLEXES CONTAINING
NUCLEIC ACIDS ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN
GENE THERAPY

(iii) NUMBER OF SEQUENCES: 43

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Connolly, Bove, Lodge, & Hutz
(B) STREET: 1220 Market Street, P.O. Box 2207
(C) CITY: Wilmington
(D) STATE: Delaware
(E) COUNTRY: U.S.A.
(F) ZIP: 19899

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk (provided in parent application)
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: Not yet assigned
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/985,526
(B) FILING DATE: 5-DEC-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: McMorrow Jr., Robert G

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (302) 658-9141
(B) TELEFAX: (302) 658-5613

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 218 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Thr Glu Glu Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro
1   5   10   15

Leu Cys Tyr His Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr
        20   25   30

Val Asp Ser Cys Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys
        35   40   45

Lys Lys Val Ser Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro
        50   55   60

Asp Gly Glu Cys Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp
65   70   75   80

Gly Trp Ser Pro Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly
        85   90   95

Asn Gly Ile Gln Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg
        100   105   110

Cys Glu Gly Ser Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys
        115   120   125

Asp Lys Arg Phe Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp
        130   135   140

Ser Ser Cys Ser Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Thr
145   150   155   160

Asn Leu Cys Ser Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly
        165   170   175

Arg Glu Ala Glu Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn

40

Gly Gly Trp Gly Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys
    195   200   205

Gly Gly Gly Val Gln Lys Arg Ser Arg Leu
   210   215

(2) INFORMATION FOR SEQ ID NO:2:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 657 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGACTGAAG AGAACAAAGA GTTGGCCAAT GAGCTGAGGC GGCCTCCCCT
ATGCTATCAC   60

AACGGAGTTC AGTACAGAAA TAACGAGGAA TGGACTGTTG ATAGCTGCAC
TGAGTGTCAC   120

TGTCAGAACT CAGTTACCAT CTGCAAAAAG GTGTCCTGCC CCATCATGCC
CTGCTCCAAT   180

GCCACAGTTC CTGATGGAGA ATGCTGTCCT CGCTGTTGGC CCAGCGACTC
TGCGGACGAT   240

GGCTGGTCTC CATGGTCCGA GTGGACCTCC TGTTCTACGA GCTGTGGCAA
TGGAATTCAG   300

CAGCGCGGCC GCTCCTGCGA TAGCCTCAAC AACCGATGTG AGGGCTCCTC
GGTCCAGACA   360

CGGACCTGCC ACATTCAGGA GTGTGACAAA AGATTTAAAC AGGATGGTGG
CTGGAGCCAC   420

TGGTCCCCGT GGTCATCTTG TTCTGTGACA TGTGGTGATG GTGTGATCAC
AAGGATCCGG   480

41

CTCTGCAACT CTCCCAGCCC CCAGATGAAT GGGAAACCCT GTGAAGGCGA
AGCGCGGGAG   540

ACCAAAGCCT GCAAGAAAGA CGCCTGCCCC ATCAATGGAG GCTGGGGTCC
TTGGTCACCA   600

TGGGACATCT GTTCTGTCAC CTGTGGAGGA GGGGTACAGA AACGTAGTCG
TCTCTAA   656

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 441 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Thr Glu Glu Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro
1   5   10   15

Leu Cys Tyr His Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr
    20   25   30

Asp Val Ser Cys Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys
    35   40   45

Lys Lys Val Ser Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro
  50   55   60

Asp Gly Glu Cys Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp
65   70   75   80

Trp Gly Ser Pro Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly
    85   90   95

Gly Asn Ile Gln Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg
    100   105   110

Cys Glu Gly Ser Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys
    115   120   125

42

Asp Lys Arg Phe Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp
    130   135   140

Ser Ser Cys Ser Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Thr
145   150   155   160

Leu Cys Asn Ser Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly
      165   170   175

Glu Ala Arg Glu Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn
      180   185   190

Gly Gly Trp Gly Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys
      195   200   205

Gly Gly Gly Val Gln Lys Arg Ser Arg Leu Cys Val Asp Ser Arg Met
    210   215   220

Thr Glu Glu Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro Leu
225   230   235   240

Cys Tyr His Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr Val
      245   250   255

Asp Ser Cys Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys Lys
      260   265   270

Lys Val Ser Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro Asp
      275   280   285

Gly Glu Cys Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp Gly
    290   295   300

Trp Ser Pro Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly Asn
305   310   315   320

Gly Ile Gln Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg Cys
      325   330   335

Glu Gly Ser Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys Asp
      340   345   350

Lys Arg Phe Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp Ser
      355   360   365

Ser Cys Ser Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Thr Leu
        370   375   380

Cys Asn Ser Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly Glu
385   390   395   400

Ala Arg Glu Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn Gly
            405   410   415

Gly Trp Gly Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys Gly
            420   425   430

Gly Gly Val Gln Lys Arg Ser Arg Leu
            435   440

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1326 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGACTGAAG AGAACAAAGA GTTGGCCAAT GAGCTGAGGC GGCCTCCCCT
ATGCTATCAC   60

AACGGAGTTC AGTACAGAAA TAACGAGGAA TGGACTGTTG ATAGCTGCAC
TGAGTGTCAC   120

TGTCAGAACT CAGTTACCAT CTGCAAAAAG GTGTCCTGCC CCATCATGCC
CTGCTCCAAT   180

GCCACAGTTC CTGATGGAGA ATGCTGTCCT CGCTGTTGGC CCAGCGACTC
TGCGGACGAT   240

GGCTGGTCTC CATGGTCCGA GTGGACCTCC TGTTCTACGA GCTGTGGCAA
TGGAATTCAG   300

CAGCGCGGCC GCTCCTGCGA TAGCCTCAAC AACCGATGTG AGGGCTCCTC

44

GGTCCAGACA   360

CGGACCTGCC ACATTCAGGA GTGTGACAAA AGATTTAAAC AGGATGGTGG
CTGGAGCCAC   420

TGGTCCCCGT GGTCATCTTG TTCTGTGACA TGTGGTGATG GTGTGATCAC
AAGGATCCGG   480

CTCTGCAACT CTCCCAGCCC CCAGATGAAT GGGAAACCCT GTGAAGGCGA
AGCGCGGGAG   540

ACCAAAGCCT GCAAGAAAGA CGCCTGCCCC ATCAATGGAG GCTGGGGTCC
TTGGTCACCA   600

TGGGACATCT GTTCTGTCAC CTGTGGAGGA GGGGTACAGA AACGTAGTCG
TCTCTGCGTC   660

GACTCTAGAA TGACTGAAGA GAACAAAGAG TTGGCCAATG AGCTGAGGCG
GCCTCCCCTA   720

TGCTATCACA ACGGAGTTCA GTACAGAAAT AACGAGGAAT GGACTGTTGA
TAGCTGCACT   780

GAGTGTCACT GTCAGAACTC AGTTACCATC TGCAAAAAGG TGTCCTGCCC
CATCATGCCC   840

TGCTCCAATG CCACAGTTCC TGATGGAGAA TGCTGTCCTC GCTGTTGGCC
CAGCGACTCT   900

GCGGACGATG GCTGGTCTCC ATGGTCCGAG TGGACCTCCT GTTCTACGAG
CTGTGGCAAT   960

GGAATTCAGC AGCGCGGCCG CTCCTGCGAT AGCCTCAACA ACCGATGTGA
GGGCTCCTCG   1020

GTCCAGACAC GGACCTGCCA CATTCAGGAG TGTGACAAAA GATTTAAACA
GGATGGTGGC   1080

TGGAGCCACT GGTCCCCGTG GTCATCTTGT TCTGTGACAT GTGGTGATGG
TGTGATCACA   1140

AGGATCCGGC TCTGCAACTC TCCCAGCCCC CAGATGAATG GGAAACCCTG
TGAAGGCGAA   1200

GCGCGGGAGA CCAAAGCCTG CAAGAAAGAC GCCTGCCCCA TCAATGGAGG

45

CTGGGGTCCT   1260

TGGTCACCAT GGGACATCTG TTCTGTCACC TGTGGAGGAG GGGTACAGAA
ACGTAGTCGT   1320

CTCTAA   1326

(2) INFORMATION FOR SEQ ID NO:5:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 6 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Tyr Ile Gly Ser Arg
1   5

(2) INFORMATION FOR SEQ ID NO:6:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 33 base pairs
     (B) TYPE: nucleic acid
     (C) STRANDEDNESS: single
     (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTCGACATGT ATATTGGTTC TCGTTAAGTC GAC   33

(2) INFORMATION FOR SEQ ID NO:7:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 25 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: linear

46

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Tyr Ile Gly Ser Arg Gly Lys Ser Tyr Ile Gly Ser Arg Gly Lys
1   5   10   15

Ser Tyr Ile Gly Ser Arg Gly Lys Ser
        20   25

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 90 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTCGACATGT ATATTGGTTC TCGTGTAAAA GTTATATTGG TTCTCGTGGT
AAAAGTTATA   60

TTGGTTCTCG TGGTAAAAGT TAAGTCGACC   90

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Leu Tyr Lys Lys Ile Ile Lys Lys Leu Leu Glu Ser
1   5   10

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 54 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTCGACATGC TTTATAAGAA GATCATCAAG AAGCTTCTTG AGAGTTAAGT CGAC
   54

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 46 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Leu Tyr Lys Lys Ile Ile Lys Lys Leu Leu Glu Ser Gly Lys Ser
1   5   10   15

Leu Tyr Lys Lys Ile Ile Lys Lys Leu Leu Glu Ser Gly Lys Ser Leu
    20   25   30

Tyr Lys Lys Ile Ile Lys Lys Leu Leu Glu Ser Gly Lys Ser
   35   40   45

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 153 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTCGACATGC TTTATAAGAA GATCATCAAG AAGCTTCTTG AGAGTGGTAA
AAGTCTTTAT   60

AAGAAGATCA TCAAGAAGCT TCTTGAGAGT GGTAAAAGTC TTTATAAGAA
GATCATCAAG   120

AAGCTTCTTG AGAGTGGTAA AAGTTAAGTC GAC   153

(2) INFORMATION FOR SEQ ID NO:13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

    Met Phe Cys Tyr Trp Lys Val Cys Trp
    1   5

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 42 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

49

GTCGACATGT TCTGTTATTG GAAGGTTTGT TGGTAAGTCG AC   42

(2) INFORMATION FOR SEQ ID NO:15:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 34 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Phe Cys Tyr Trp Lys Val Cys Trp Gly Lys Ser Phe Cys Tyr Trp
1   5   10   15

Lys Val Cys Trp Gly Lys Ser Phe Cys Tyr Trp Lys Val Cys Trp Gly
     20   25   30

Lys Ser

(2) INFORMATION FOR SEQ ID NO:16:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 117 base pairs
     (B) TYPE: nucleic acid
     (C) STRANDEDNESS: single
     (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCGACATGT TCTGTTATTG GAAGGTTTGT TGGGGTAAAA GTTTCTGTTA
TTGGAAGGTT   60

TGTTGGGGTA AAAGTTTCTG TTATTGGAAG GTTTGTTGGG GTAAAAGTTA
AGTCGAC   117

(2) INFORMATION FOR SEQ ID NO:17:

50

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 5 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Gly Arg Gly Asp
1   5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 30 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GTCGACATGG GTCGTGGTGA TTAAGTCGAC   30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 22 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Gly Arg Gly Asp Gly Lys Ser Gly Arg Gly Asp Gly Lys Ser Gly
1   5   10   15

51

Arg Gly Asp Gly Lys Ser
            20

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 81 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTCGACATGG GTCGTGGTGA TGGTAAAAGT GGTCGTGGTG ATGGTAAAAG
TGGTCGTGGT   60

GATGGTAAAA GTTAAGTCGA C   81

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 210 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Val Tyr Leu Ser Glu Cys Lys Thr Gly Ile Gly Asn Gly Tyr Arg
1   5   10   15

Gly Thr Met Ser Arg Thr Lys Ser Gly Val Ala Cys Gln Lys Trp Gly
      20   25   30

Ala Thr Phe Pro His Val Pro Asn Tyr Ser Pro Ser Thr His Pro Asn
     35   40   45

Glu Gly Leu Glu Glu Asn Tyr Cys Arg Asn Pro Asp Asn Asp Glu Gln

52

Gly Pro Trp Cys Tyr Thr Thr Asp Pro Asp Lys Arg Tyr Asp Tyr Cys
65   70   75   80

Asn Ile Pro Glu Cys Glu Glu Glu Cys Met Tyr Cys Ser Gly Glu Lys
        85   90   95

Tyr Glu Gly Lys Ile Ser Lys Thr Met Ser Gly Lys Asp Cys Gln Ala
        100   105   110

Trp Asp Ser Gln Ser Pro His Ala His Gly Tyr Ile Pro Ala Lys Phe
        115   120   125

Pro Ser Lys Asn Leu Lys Met Asn Tyr Cys His Asn Pro Asp Gly Glu
        130   135   140

Pro Arg Pro Trp Cys Phe Thr Thr Asp Pro Thr Lys Arg Trp Glu Tyr
145   150   155   160

Cys Asp Ile Pro Arg Cys Thr Thr Pro Pro Pro Pro Ser Pro Thr
        165   170   175

Tyr Gln Cys Leu Lys Gly Arg Gly Glu Asn Tyr Arg Gly Thr Val Ser
        180   185   190

Val Thr Val Ser Gly Lys Thr Cys Gln Arg Trp Ser Glu Gln Thr Pro
        195   200   205

His Arg
        210

(2) INFORMATION FOR SEQ ID NO:22:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 645 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTCGACATGG TGTATCTGTC AGAATGTAAG ACCGGCATCG GCAACGGCTA
CAGAGGAACC   60

ATGTCCAGGA CAAAGAGTGG TGTTGCCTGT CAAAAGTGGG GTGCCACGTT
CCCCCACGTA   120

CCCAACTACT CTCCCAGTAC ACATCCCAAT GAGGGACTAG AAGAGAACTA
CTGTAGGAAC   180

CCAGACAATG ATGAACAAGG GCCTTGGTGC TACACTACAG ATCCGGACAA
GAGATATGAC   240

TACTGCAACA TTCCTGAATG TGAAGAGGAA TGCATGTACT GCAGTGGAGA
AAAGTATGAG   300

GGCAAAATCT CCAAGACCAT GTCTGGACTT GACTGCCAGG CCTGGGATTC
TCAGAGCCCA   360

CATGCTCATG GATACATCCC TGCCAAATTT CCAAGCAAGA ACCTGAAGAT
GAATTATTGC   420

CACAACCCTG ACGGGGAGCC AAGGCCCTGG TGCTTCACAA CAGACCCCAC
CAAACGCTGG   480

GAATACTGTG ACATCCCCCG CTGCACAACA CCCCCGCCCC CACCCAGCCC
AACCTACCAA   540

TGTCTGAAAG GAAGAGGTGA AAATTACCGA GGGACCGTGT CTGTCACCGT
GTCTGGGAAA   600

ACCTGTCAGC GCTGGAGTGA GCAAACCCCT CATAGGTGAG TCGAC   644

(2) INFORMATION FOR SEQ ID NO:23:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 423 amino acids
     (B) TYPE: amino acid
     (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

54

Met Val Tyr Leu Ser Glu Cys Lys Thr Gly Ile Gly Asn Gly Tyr Arg
1   5   10   15

Gly Thr Met Ser Arg Thr Lys Ser Gly Val Ala Cys Gln Lys Trp Gly
        20   25   30

Ala Thr Phe Pro His Val Pro Asn Tyr Ser Pro Ser Thr His Pro Asn
        35   40   45

Glu Gly Leu Glu Glu Asn Tyr Cys Arg Asn Pro Asp Asn Asp Glu Gln
        50   55   60

Gly Pro Trp Cys Tyr Thr Thr Asp Pro Asp Lys Arg Tyr Asp Tyr Cys
65   70   75   80

Asn Ile Pro Glu Cys Glu Glu Glu Cys Met Tyr Cys Ser Gly Glu Lys
            85   90   95

Tyr Glu Gly Lys Ile Ser Lys Thr Met Ser Gly Lys Asp Cys Gln Ala
            100   105   110

Trp Asp Ser Gln Ser Pro His Ala His Gly Tyr Ile Pro Ala Lys Phe
            115   120   125

Pro Ser Lys Asn Leu Lys Met Asn Tyr Cys His Asn Pro Asp Gly Glu
            130   135   140

Pro Arg Pro Trp Cys Phe Thr Thr Asp Pro Thr Lys Arg Trp Glu Tyr
145   150   155   160

Cys Asp Ile Pro Arg Cys Thr Thr Pro Pro Pro Pro Ser Pro Thr
            165   170   175

Tyr Gln Cys Leu Lys Gly Arg Gly Glu Asn Tyr Arg Gly Thr Val Ser
            180   185   190

Val Thr Val Ser Gly Lys Thr Cys Gln Arg Trp Ser Glu Gln Thr Pro
            195   200   205

His Arg Gly Lys Ser Met Val Tyr Leu Ser Glu Cys Lys Thr Gly Ile
            210   215   220

Gly Asn Gly Tyr Arg Gly Thr Met Ser Arg Thr Lys Ser Gly Val Ala
225   230   235   240

55

Cys Gln Lys Trp Gly Ala Thr Phe Pro His Val Pro Asn Tyr Ser Pro
        245   250   255

Ser Thr His Pro Asn Glu Gly Leu Glu Glu Asn Tyr Cys Arg Asn Pro
        260   265   270

Asp Asn Asp Glu Gln Gly Pro Trp Cys Tyr Thr Thr Asp Pro Asp Lys
        275   280   285

Arg Tyr Asp Tyr Cys Asn Ile Pro Glu Cys Glu Glu Glu Cys Met Tyr
       290   295   300

Cys Ser Gly Glu Lys Tyr Glu Gly Lys Ile Ser Lys Thr Met Ser Gly
305   310   315   320

Lys Asp Cys Gln Ala Trp Asp Ser Gln Ser Pro His Ala His Gly Tyr
        325   330   335

Ile Pro Ala Lys Phe Pro Ser Lys Asn Leu Lys Met Asn Tyr Cys His
        340   345   350

Asn Pro Asp Gly Glu Pro Arg Pro Trp Cys Phe Thr Thr Asp Pro Thr
        355   360   365

Lys Arg Trp Glu Tyr Cys Asp Ile Pro Arg Cys Thr Thr Pro Pro Pro
       370   375   380

Pro Pro Ser Pro Thr Tyr Gln Cys Leu Lys Gly Arg Gly Glu Asn Tyr
385   390   395   400

Arg Gly Thr Val Ser Val Thr Val Ser Gly Lys Thr Cys Gln Arg Trp
        405   410   415

Ser Glu Gln Thr Pro His Arg
        420

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1284 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTCGACATGG TGTATCTGTC AGAATGTAAG ACCGGCATCG GCAACGGCTA
CAGAGGAACC   60

ATGTCCAGGA CAAAGAGTGG TGTTGCCTGT CAAAAGTGGG GTGCCACGTT
CCCCCACGTA   120

CCCAACTACT CTCCCAGTAC ACATCCCAAT GAGGGACTAG AAGAGAACTA
CTGTAGGAAC   180

CCAGACAATG ATGAACAAGG GCCTTGGTGC TACACTACAG ATCCGGACAA
GAGATATGAC   240

TACTGCAACA TTCCTGAATG TGAAGAGGAA TGCATGTACT GCAGTGGAGA
AAAGTATGAG   300

GGCAAAATCT CCAAGACCAT GTCTGGACTT GACTGCCAGG CCTGGGATTC
TCAGAGCCCA   360

CATGCTCATG GATACATCCC TGCCAAATTT CCAAGCAAGA ACCTGAAGAT
GAATTATTGC   420

CACAACCCTG ACGGGGAGCC AAGGCCCTGG TGCTTCACAA CAGACCCCAC
CAAACGCTGG   480

GAATACTGTG ACATCCCCCG CTGCACAACA CCCCCGCCCC CACCCAGCCC
AACCTACCAA   540

TGTCTGAAAG GAAGAGGTGA AAATTACCGA GGGACCGTGT CTGTCACCGT
GTCTGGGAAA   600

ACCTGTCAGC GCTGGAGTGA GCAAACCCCT CATAGGGGTA AAAGTATGGT
GTATCTGTCA   660

GAATGTAAGA CCGGCATCGG CAACGGCTAC AGAGGAACCA TGTCCAGGAC
AAAGAGTGGT   720

GTTGCCTGTC AAAAGTGGGG TGCCACGTTC CCCCACGTAC CCAACTACTC
TCCCAGTACA   780

CATCCCAATG AGGGACTAGA AGAGAACTAC TGTAGGAACC CAGACAATGA
TGAACAAGGG   840

57

CCTTGGTGCT ACACTACAGA TCCGGACAAG AGATATGACT ACTGCAACAT
TCCTGAATGT   900

GAAGAGGAAT GCATGTACTG CAGTGGAGAA AAGTATGAGG GCAAAATCTC
CAAGACCATG   960

TCTGGACTTG ACTGCCAGGC CTGGGATTCT CAGAGCCCAC ATGCTCATGG
ATACATCCCT   1020

GCCAAATTTC CAAGCAAGAA CCTGAAGATG AATTATTGCC ACAACCCTGA
CGGGGAGCCA   1080

AGGCCCTGGT GCTTCACAAC AGACCCCACC AAACGCTGGG AATACTGTGA
CATCCCCCGC   1140

TGCACAACAC CCCCGCCCCC ACCCAGCCCA ACCTACCAAT GTCTGAAAGG
AAGAGGTGAA   1200

AATTACCGAG GGACCGTGTC TGTCACCGTG TCTGGGAAAA CCTGTCAGCG
CTGGAGTGAG   1260

CAAACCCCTC ATAGGTGAGT CGAC   1284

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 125 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

    Met Leu Pro Ile Cys Pro Gly Gly Ala Ala Arg Cys Gln Val Thr Leu
    1   5   10   15

    Arg Glu Leu Phe Asp Arg Ala Val Val Leu Ser His Tyr Ile His Asn
            20   25   30

    Leu Ser Ser Glu Met Phe Ser Glu Phe Glu Lys Arg Tyr Thr His Gly
            35   40   45

Arg Gly Phe Ile Thr Lys Ala Ile Asn Ser Cys His Thr Ser Ser Leu
   50   55   60

Ala Thr Pro Glu Asp Lys Glu Gln Ala Gln Gln Met Asn Gln Lys Asp
65   70   75   80

Phe Leu Ser Leu Ile Val Ser Ile Leu Arg Ser Trp Asn Glu Pro Leu
            85   90   95

Tyr His Leu Val Thr Glu Val Arg Gly Met Gln Glu Ala Pro Gln Ala
            100   105   110

Ile Leu Ser Lys Ala Val Glu Ile Glu Glu Gln Thr Lys
            115   120   125

(2) INFORMATION FOR SEQ ID NO:26:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 390 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GTCGACATGT TGCCCATCTG TCCCGGCGGG GCTGCCCGAT GCCAGGTGAC
CCTTCGAGAC   60

CTGTTTGACC GCGCCGTCGT CCTGTCCCAC TACATCCATA ACCTCTCCTC
AGAAATGTTC   120

AGCGAATTCG ATAAACGGTA TACCCATGGC CGGGGGTTCA TTACCAAGGC
CATCAACAGC   180

TGCCACACTT CTTCCCTTGC CACCCCCGAA GACAAGGAGC AAGCCCAACA
GATGAATCAA   240

AAAGACTTTC TGAGCCTGAT AGTCAGCATA TTGCGATCCT GGAATGAGCC
TCTGTATCAT   300

CTGGTCACGG AAGTACGTGG TATGCAAGAA GCCCCGGAGG CTATCCTATC

59

CAAAGCTGTA   360

GAGATTGAGG AGCAAACCAA ATAAGTCGAC   390

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 253 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Leu Pro Ile Cys Pro Gly Gly Ala Ala Arg Cys Gln Val Thr Leu
1   5   10   15

Arg Glu Leu Phe Asp Arg Ala Val Val Leu Ser His Tyr Ile His Asn
     20   25   30

Leu Ser Ser Glu Met Phe Ser Glu Phe Glu Lys Arg Tyr Thr His Gly
    35   40   45

Arg Gly Phe Ile Thr Lys Ala Ile Asn Ser Cys His Thr Ser Ser Leu
   50   55   60

Ala Thr Pro Glu Asp Lys Glu Gln Ala Gln Gln Met Asn Gln Lys Asp
65   70   75   80

Phe Leu Ser Leu Ile Val Ser Ile Leu Arg Ser Trp Asn Glu Pro Leu
     85   90   95

Tyr His Leu Val Thr Glu Val Arg Gly Met Gln Glu Ala Pro Gln Ala
    100   105   110

Ile Leu Ser Lys Ala Val Glu Ile Glu Glu Gln Thr Lys Gly Lys Ser
    115   120   125

Met Leu Pro Ile Cys Pro Gly Gly Ala Ala Arg Cys Gln Val Thr Leu
   130   135   140

Arg Glu Leu Phe Asp Arg Ala Val Val Leu Ser His Tyr Ile His Asn

60

Leu Ser Ser Glu Met Phe Ser Glu Phe Glu Lys Arg Tyr Thr His Gly
         165   170   175

Arg Gly Phe Ile Thr Lys Ala Ile Asn Ser Cys His Thr Ser Ser Leu
        180   185   190

Ala Thr Pro Glu Asp Lys Glu Gln Ala Gln Gln Met Asn Gln Lys Asp
      195   200   205

Phe Leu Ser Leu Ile Val Ser Ile Leu Arg Ser Trp Asn Glu Pro Leu
210   215   220

Tyr His Leu Val Thr Glu Val Arg Gly Met Gln Glu Ala Pro Gln Ala
225   230   235   240

Ile Leu Ser Lys Ala Val Glu Ile Glu Glu Gln Thr Lys
        245   250

(2) INFORMATION FOR SEQ ID NO:28:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 771 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCGACATGT TGCCCATCTG TCCCGGCGGG GCTGCCCGAT GCCAGGTGAC
CCTTCGAGAC   60

CTGTTTGACC GCGCCGTCGT CCTGTCCCAC TACATCCATA ACCTCTCCTC
AGAAATGTTC   120

AGCGAATTCG ATAAACGGTA TACCCATGGC CGGGGGTTCA TTACCAAGGC
CATCAACAGC   180

TGCCACACTT CTTCCCTTGC CACCCCCGAA GACAAGGAGC AAGCCCAACA
GATGAATCAA   240

AAAGACTTTC TGAGCCTGAT AGTCAGCATA TTGCGATCCT GGAATGAGCC
TCTGTATCAT   300

CTGGTCACGG AAGTACGTGG TATGCAAGAA GCCCCGGAGG CTATCCTATC
CAAAGCTGTA   360

GAGATTGAGG AGCAAACCGG TAAAAGTATG TTGCCCATCT GTCCCGGCGG
GGCTGCCCGA   420

TGCCAGGTGA CCCTTCGAGA CCTGTTTGAC CGCGCCGTCG TCCTGTCCCA
CTACATCCAT   480

AACCTCTCCT CAGAAATGTT CAGCGAATTC GATAAACGGT ATACCCATGG
CCGGGGGTTC   540

ATTACCAAGG CCATCAACAG CTGCCACACT TCTTCCCTTG CCACCCCCGA
AGACAAGGAG   600

CAAGCCCAAC AGATGAATCA AAAAGACTTT CTGAGCCTGA TAGTCAGCAT
ATTGCGATCC   660

TGGAATGAGC CTCTGTATCA TCTGGTCACG GAAGTACGTG GTATGCAAGA
AGCCCCGGAG   720

GCTATCCTAT CCAAAGCTGT AGAGATTGAG GAGCAAACCA AATAAGTCGA C
 771

(2) INFORMATION FOR SEQ ID NO:29:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 161 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATGCTGAGGC GGCCTCCCCT ATGCTATCAC AACGGAGTTC AGTACAGAAA
TAACGGTAAA   60

AGATCCCCGT GGTCATCTTG TTCTGTGACA TGTGGTGATG GTGTGATGGT

62

AAAAGAAGTG   120

GTACCCTGTA GACAAGACAG TGGACACCTC CTCCCCATTA A   161

(2) INFORMATION FOR SEQ ID NO:30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 63 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Leu Arg Arg Pro Pro Leu Cys Tyr His Asn Gly Val Gln Tyr Arg
1   5   10   15

Asn Asn Glu Glu Trp Thr Val Asp Ser Gly Lys Ser Ser Pro Trp Ser
          20   25   30

Ser Cys Ser Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Gly Lys
          35   40   45

Ser Ser Pro Trp Asp Ile Cys Ser Val Thr Cys Gly Gly Gly Val
          50   55   60

(2) INFORMATION FOR SEQ ID NO:31:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 185 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ATGCTGAGGC GGCCTCCCCT ATGCTATCAC AACGGAGTTC AGTACAGAAA
TAACGGTAAA   60

63

AGATCCCCGT GGTCATCTTG TTCTGTGACA TGTGGTGATG GTGTGATGGT
AAAAGAAGTG   120

GTACCCTGTA GACAAGACAG TGGACACCTC CTCCCCATTA TATTGGTTCT
CGTGGTAAAA   180

GATAA   185

(2) INFORMATION FOR SEQ ID NO:32:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TAGGTCTAGA ATGACTGAAG AGAACAAAGA G   31

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATGGTCTAGA TTAGAGACGA CTACGTTTCT G   31

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 805 amino acids
        (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Glu Ser Lys Ala Leu Leu Ala Val Ala Leu Trp Phe Cys Val Glu
1   5   10   15

Thr Arg Ala Ala Ser Val Gly Leu Pro Gly Asp Phe Leu His Pro Pro
        20   25   30

Lys Leu Ser Thr Gln Lys Asp Ile Leu Thr Ile Leu Ala Asn Thr Thr
        35   40   45

Leu Gln Ile Thr Cys Arg Gly Gln Arg Asp Leu Asp Trp Leu Trp Pro
        50   55   60

Asn Ala Gln Arg Asp Ser Glu Glu Arg Val Leu Val Thr Glu Cys Gly
        65   70   75   80

Gly Gly Asp Ser Ile Phe Cys Lys Thr Leu Thr Ile Pro Arg Val Val
            85   90   95

Gly Asn Asp Thr Gly Ala Tyr Lys Cys Ser Tyr Arg Asp Val Asp Ile
            100   105   110

Ala Ser Thr Val Tyr Val Tyr Val Arg Asp Tyr Arg Ser Pro Phe Ile
            115   120   125

Ala Ser Val Ser Asp Gln His Gly Ile Val Tyr Ile Thr Glu Asn Lys
            130   135   140

Asn Lys Thr Val Val Ile Pro Cys Arg Gly Ser Ile Ser Asn Leu Asn
        145   150   155   160

Val Ser Leu Cys Ala Arg Tyr Pro Glu Lys Arg Phe Val Pro Asp Gly
            165   170   175

Asn Arg Ile Ser Trp Asp Ser Glu Ile Gly Phe Thr Leu Pro Ser Tyr
            180   185   190

Met Ile Ser Tyr Ala Gly Met Val Phe Cys Glu Ala Lys Ile Asn Asp

65

Glu Thr Tyr Gln Ser Ile Met Tyr Ile Val Val Val Val Gly Tyr Arg
    210   215   220

Ile Tyr Asp Val Ile Leu Ser Pro Pro His Glu Ile Glu Leu Ser Ala
225   230   235   240

Gly Glu Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val
          245   250   255

Gly Leu Asp Phe Thr Trp His Ser Pro Pro Ser Lys Ser His His Lys
          260   265   270

Lys Ile Val Asn Arg Asp Val Lys Pro Phe Pro Gly Thr Val Ala Lys

          275   280   285

Met Phe Lys Ser Thr Leu Thr Ile Glu Ser Val Thr Lys Ser Asp Gln
    290   295   300

Gly Glu Tyr Thr Cys Val Ala Ser Ser Gly Arg Met Ile Lys Arg Asn
305   310   315   320

Arg Thr Phe Val Arg Val His Thr Lys Pro Phe Ile Ala Phe Gly Ser
          325   330   335

Gly Met Lys Ser Leu Val Glu Ala Thr Val Gly Ser Gln Val Arg Ile
          340   345   350

Pro Val Lys Tyr Leu Ser Tyr Pro Ala Pro Asp Ile Lys Trp Tyr Arg
    355   360   365

Asn Gly Arg Pro Ile Glu Ser Asn Tyr Thr Met Ile Val Gly Asp Glu
    370   375   380

Leu Thr Ile Met Glu Val Thr Glu Arg Asp Ala Gly Asn Tyr Thr Val
385   390   395   400

Ile Leu Thr Asn Pro Ile Ser Met Glu Lys Gln Ser His Met Val Ser
          405   410   415

Leu Val Val Asn Val Pro Pro Gln Ile Gly Glu Lys Ala Leu Ile Ser
    420   425   430

66

Pro Met Asp Ser Tyr Gly Tyr Gly Thr Met Gln Thr Leu Thr Cys Thr
    435   440   445

Val Tyr Ala Asn Pro Pro Leu His His Ile Gln Trp Tyr Trp Gln Leu
    450   455   460

Glu Glu Ala Cys Ser Tyr Arg Pro Gly Gln Thr Ser Pro Tyr Ala Cys
465   470   475   480

Lys Glu Trp Arg His Val Glu Asp Phe Gln Gly Gly Asn Lys Ile Glu
    485   490   495

Val Thr Lys Asn Gln Tyr Ala Leu Ile Glu Gly Lys Asn Lys Thr Val
    500   505   510

Ser Thr Leu Val Ile Gln Ala Ala Asn Val Ser Ala Leu Tyr Lys Cys
    515   520   525

Glu Ala Ile Asn Lys Ala Gly Arg Gly Glu Arg Val Ile Ser Phe His

   530   535   540

Val Ile Arg Gly Pro Glu Ile Thr Val Gln Pro Ala Ala Gln Pro Thr
545   550   555   560

Glu Gln Glu Ser Val Ser Leu Leu Cys Thr Ala Asp Arg Asn Thr Phe
    565   570   575

Glu Asn Leu Thr Trp Tyr Lys Leu Gly Ser Gln Ala Thr Ser Val His
    580   585   590

Met Gly Glu Ser Leu Thr Pro Val Cys Lys Asn Leu Asp Ala Leu Trp
    595   600   605

Lys Leu Asn Gly Thr Met Phe Ser Asn Ser Thr Asn Asp Ile Leu Ile
    610   615   620

Val Ala Phe Gln Asn Ala Ser Leu Gln Asp Gln Gly Asp Tyr Val Cys
625   630   635   640

Ser Ala Gln Asp Lys Lys Thr Lys Lys Arg His Cys Leu Val Lys Gln
    645   650   655

Leu Ile Ile Leu Glu Arg Met Ala Pro Met Ile Thr Gly Asn Leu Glu
    660   665   670

67

Asn Gln Thr Thr Thr Ile Gly Glu Thr Ile Glu Val Thr Cys Pro Ala
675   680   685

Ser Gly Asn Pro Thr Pro His Ile Thr Trp Phe Lys Asp Asn Glu Thr
690   695   700

Leu Val Glu Asp Ser Gly Ile Val Leu Arg Asp Gly Asn Arg Asn Leu
705   710   715   720

Thr Ile Arg Arg Val Arg Lys Glu Asp Gly Gly Leu Tyr Thr Cys Gln
725   730   735

Ala Cys Asn Val Leu Gly Cys Ala Arg Ala Glu Thr Leu Phe Ile Ile
740   745   750

Glu Gly Ala Gln Glu Lys Thr Asn Leu Glu Val Ile Ile Leu Val Gly
755   760   765

Thr Ala Val Ile Ala Met Phe Phe Trp Leu Leu Leu Val Ile Leu Val
770   775   780

Arg Thr Val Lys Arg Ala Asn Glu Gly Glu Leu Lys Thr Gly Tyr Leu
785   790   795   800

Ser Ile Val Met Asp
805

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2431 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AGACGTCATG GAGAGCAAGG CGCTGCTAGC TGTCGCTCTG TGGTTCTGCG
TGGAGACCCG   60

AGCCGCCTCT GTGGGTTTGC CTGGCGATTT TCTCCATCCC CCCAAGCTCA
GCACACAGAA   120

AGACATACTG ACAATTTTGG CAAATACAAC CCTTCAGATT ACTTGCAGGG
GACAGCGGGA   180

CCTGGACTGG CTTTGGCCCA ATGCTCAGCG TGATTCTGAG GAAAGGGTAT
TGGTGACTGA   240

ATGCGGCGGT GGTGACAGTA TCTTCTGCAA AACACTCACC ATTCCCAGGG
TGGTTGGAAA   300

TGATACTGGA GCCTACAAGT GCTCGTACCG GGACGTCGAC ATAGCCTCCA
CTGTTTATGT   360

CTATGTTCGA GATTACAGAT CACCATTCAT CGCCTCTGTC AGTGACCAGC
ATGGCATCGT   420

GTACATCACC GAGAACAAGA ACAAAACTGT GGTGATCCCC TGCCGAGGGT
CGATTTCAAA   480

CCTCAATGTG TCTCTTTGCG CTAGGTATCC AGAAAAGAGA TTTGTTCCGG
ATGGAAACAG   540

AATTTCCTGG GACAGCGAGA TAGGCTTTAC TCTCCCCAGT TACATGATCA
GCTATGCCGG   600

CATGGTCTTC TGTGAGGCAA AGATCAATGA TGAAACCTAT CAGTCTATCA
TGTACATAGT   660

TGTGGTTGTA GGATATAGGA TTTATGATGT GATTCTGAGC CCCCCGCATG
AAATTGAGCT   720

ATCTGCCGGA GAAAAACTTG TCTTAAATTG TACAGCGAGA ACAGAGCTCA
ATGTGGGGCT   780

TGATTTCACC TGGCACTCTC CACCTTCAAA GTCTCATCAT AAGAAGATTG
TAAACCGGGA   840

TGTGAAACCC TTTCCTGGGA CTGTGGCGAA GATGTTTTTG AGCACCTTGA
CAATAGAAAG   900

TGTGACCAAG AGTGACCAAG GGGAATACAC CTGTGTAGCG TCCAGTGGAC
GGATGATCAA   960

69

GAGAAATAGA ACATTTGTCC GAGTTCACAC AAAGCCTTTT ATTGCTTTCG
GTAGTGGGAT   1020

GAAATCTTTG GTGGAAGCCA CAGTGGGCAG TCAAGTCCGA ATCCCTGTGA
AGTATCTCAG   1080

TTACCCAGCT CCTGATATCA AATGGTACAG AAATGGAAGG CCCATTGAGT
CCAACTACAC   1140

AATGATTGTT GGCGATGAAC TCACCATCAT GGAAGTGACT GAAAGAGATG
CAGGAAACTA   1200

CACGGTCATC CTCACCAACC CCATTTCAAT GGAGAAACAG AGCCACATGG
TCTCTCTGGT   1260

TGTGAATGTC CCACCCCAGA TCGGTGAGAA AGCCTTGATC TCGCCTATGG
ATTCCTACCA   1320

GTATGGGACC ATGCAGACAT TGACATGCAC AGTCTACGCC AACCCTCCCC
TGCACCACAT   1380

CCAGTGGTAC TGGCAGCTAG AAGAAGCCTG CTCCTACAGA CCCGGCCAAA
CAAGCCCGTA   1440

TGCTTGTAAA GAATGGAGAC ACGTGGAGGA TTTCCAGGGG GGAAACAAGA
TCGAAGTCAC   1500

CAAAAACCAA TATGCCCTGA TTGAAGGAAA AAACAAAACT GTAAGTACGC
TGGTCATCCA   1560

AGCTGCCAAC GTGTCAGCGT TGTACAAATG TGAAGCCATC AACAAAGCGG
GACGAGGAGA   1620

GAGGGTCATC TCCTTCCATG TGATCAGGGG TCCTGAAATT ACTGTGCAAC
CTGCTGCCCA   1680

GCCAACTGAG CAGGAGAGTG TGTCCCTGTT GTGCACTGCA GACAGAAATA
CGTTTGAGAA   1740

CCTCACGTGG TACAAGCTTG GCTCACAGGC AACATCGGTC CACATGGGCG
AATCACTCAC   1800

ACCAGTTTGC AAGAACTTGG ATGCTCTTTG GAAACTGAAT GGCACCATGT
TTTCTAACAG   1860

70

CACAAATGAC ATCTTGATTG TGGCATTTCA GAATGCCTCT CTGCAGGACC
AAGGCGACTA   1920

TGTTTGCTCT GCTCAAGATA AGAAGACCAA GAAAAGACAT TGCCTGGTCA
AACAGCTCAT   1980

CATCCTAGAG CGCATGGCAC CCATGATCAC CGGAAATCTG GAGAATCAGA
CAACAACCAT   2040

TGGCGAGACC ATTGAAGTGA CTTGCCCAGC ATCTGGAAAT CCTACCCCAC
ACATTACATG   2100

GTTCAAAGAC AACGAGACCC TGGTAGAAGA TTCAGGCATT GTACTGAGAG
ATGGGAACCG   2160

GAACCTGACT ATCCGCAGGG TGAGGAAGGA GGATGGAGGC CTCTACACCT
GCCAGGCCTG   2220

CAATGTCCTT GGCTGTGCAA GAGCGGAGAC GCTCTTCATA ATAGAAGGTG
CCCAGGAAAA   2280

GACCAACTTG GAAGTCATTA TCCTCGTCGG CACTGCAGTG ATTGCCATGT
TCTTCTGGCT   2340

CCTTCTTGTC ATTCTCGTAC GGACCGTTAA GCGGGCCAAT GAAGGGGAAC
TGAAGACAGG   2400

CTACTTGTCT ATTGTCATGG ATTAAGACGT C   2431

(2) INFORMATION FOR SEQ ID NO:36:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 185 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

    Met His Thr His Gln Asp Phe Gln Pro Val Leu His Leu Val Ala Leu
    1   5   10   15

71

Asn Thr Pro Leu Ser Gly Gly Met Arg Gly Ile Arg Gly Ala Asp Phe
      20   25   30

Gln Cys Phe Asn Asn Ala Arg Val Gly Leu Ser Gly Thr Phe Arg Ala
      35   40   45

Phe Leu Ser Ser Arg Leu Gln Asp Leu Tyr Ser Ile Val Arg Arg Ala
    50   55   60

Asp Arg Gly Ser Val Pro Ile Val Gln Asn Leu Arg Asp Glu Val Leu
65   70   75   80

Ser Pro Ser Trp Asp Ser Leu Phe Ser Gly Ser Gln Gly Gln Leu Gln
      85   90   95

Pro Gly Ala Arg Ile Phe Ser Phe Asp Gly Arg Asp Val Leu Arg His
     100   105   110

Pro Ala Trp Pro Gln Arg Ser Val Trp His Gly Ser Asp Pro Ser Gly
     115   120   125

Arg Arg Leu Met Glu Ser Tyr Cys Glu Thr Trp Arg Thr Glu Thr Thr
    130   135   140

Gly Ala Thr Gly Gln Ala Ser Ser Leu Leu Ser Gly Arg Leu Leu Glu
145   150   155   160

Gln Arg Ala Ala Ser Cys His Asp Ser Tyr Ile Val Leu Cys Ile Glu
     165   170   175

Asn Ser Phe Met Thr Ser Phe Ser Arg
    180   185

(2) INFORMATION FOR SEQ ID NO:37:

  (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 565 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AGACGTCATG CATACTCATC AGGACTTTCA GCCAGTGCTC CACCTGGTGG
CACTGAACAC   60

CCCCCTGTCT GGAGGCATGC GTGGTATCCG TGGAGCAGAT TTCCAGTGCT
TCCAGCAAGC   120

CCGAGCCGTG GGGCTGTCGG GCACCTTCCG GGCTTTCCTG TCCTCTAGGC
TGCAGGATCT   180

CTATAGCATC GTGCGCCGTG CTGACCGGGG GTCTGTGCCC ATCGTCAACC
TGAAGGACGA   240

GGTGCTATCT CCCAGCTGGG ACTCCCTGTT TTCTGGCTCC CAGGGTCAAC
TGCAACCCGG   300

GGCCCGCATC TTTTCTTTTG ACGGCAGAGA TGTCCTGAGA CACCCAGCCT
GGCCGCAGAA   360

GAGCGTATGG CACGGCTCGG ACCCCAGTGG GCGGAGGCTG ATGGAGAGTT
ACTGTGAGAC   420

ATGGCGAACT GAAACTACTG GGGCTACAGG TCAGGCCTCC TCCCTGCTGT
CAGGCAGGCT   480

CCTGGAACAG AAAGCTGCGA GCTGCCACAA CAGCTACATC GTCCTGTGCA
TTGAGAATAG   540

CTTCATGACC TCTTTCTCCA AATAG   565

(2) INFORMATION FOR SEQ ID NO:38:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 43 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

73

CTATCGTCGA CATGTATATT GGTTCTCGTT AAGTCGACCT ATC 43

(2) INFORMATION FOR SEQ ID NO:39:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 43 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GATAGGTCGA CTTAACGAGA ACCAATATAC ATGTCGACGA TAG 43

(2) INFORMATION FOR SEQ ID NO:40:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

AGTATCTAGA ATGAGTGTAT CTGTCACAAT G 31

(2) INFORMATION FOR SEQ ID NO:41:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAATTCTAGA TCACCTATGA GGGGTTTGCT C   31

(2) INFORMATION FOR SEQ ID NO:42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 93 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTATCGTCGA CATGTATATT GGTTCTCGTA AAAGATATAT TGGTTCTCGT
GGTAAAAGAG   60

ATGGTTCTCG TGGTAAAAGA TAAGTGACCT ATC   93

(2) INFORMATION FOR SEQ ID NO:43:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 15 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GATAGGTCGA CTTAT   15

76814